PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Molecular dynamics study of HIV-1 RT-DNA-nevirapine complexes explains NNRTI inhibition and resistance by connection mutations

R. S. K. Vijayan,[1,2] Eddy Arnold,[1,2] and Kalyan Das[1,2]*

[1] Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, New Jersey 08854

[2] Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854

ABSTRACT

HIV-1 reverse transcriptase (RT) is a multifunctional enzyme that is targeted by nucleoside analogs (NRTIs) and non-nucleoside RT inhibitors (NNRTIs). NNRTIs are allosteric inhibitors of RT, and constitute an integral part of several highly active antiretroviral therapy regimens. Under selective pressure, HIV-1 acquires resistance against NNRTIs primarily by selecting mutations around the NNRTI pocket. Complete RT sequencing of clinical isolates revealed that spatially distal mutations arising in connection and the RNase H domain also confer NNRTI resistance and contribute to NRTI resistance. However, the precise structural mechanism by which the connection domain mutations confer NNRTI resistance is poorly understood. We performed 50-ns molecular dynamics (MD) simulations, followed by essential dynamics, free-energy landscape analyses, and network analyses of RT-DNA, RT-DNA-nevirapine (NVP), and N348I/T369I mutant RT-DNA-NVP complexes. MD simulation studies revealed altered global motions and restricted conformational landscape of RT upon NVP binding. Analysis of protein structure network parameters demonstrated a dissortative hub pattern in the RT-DNA complex and an assortative hub pattern in the RT-DNA-NVP complex suggesting enhanced rigidity of RT upon NVP binding. The connection subdomain mutations N348I/T369I did not induce any significant structural change; rather, these mutations modulate the conformational dynamics and alter the long-range allosteric communication network between the connection subdomain and NNRTI pocket. Insights from the present study provide a structural basis for the biochemical and clinical findings on drug resistance caused by the connection and RNase H mutations.

Proteins 2014; 82:815–829.

Key words: drug resistance; connection mutations; N348I; T369I; RNase H; protein structural network; allosteric communication.

INTRODUCTION

HIV-1 reverse transcriptase (RT) catalyzes the copying of the viral single-stranded (ss) RNA genome into double-stranded DNA (dsDNA). RT is an asymmetric heterodimer consisting of two subunits of molecular mass 66 (p66) and 51 kDa (p51).[1,2] The p66 subunit contains a DNA polymerase domain and a ribonuclease H (RNase H) domain. The polymerase activity is responsible for DNA synthesis complementing an RNA or DNA template, and the RNase H activity cleaves the RNA strand of an RNA/DNA duplex. The polymerase domain anthropomorphically mimics a right hand that consists of fingers, palm, thumb, and connection subdomains.[3] The p51 subunit is derived by proteolytic cleavage of the C-terminal RNase H domain from a p66 chain. The subdomains in the p51 subunit are arranged differently compared to the organization in p66, and p51 subunit largely plays a structural role. Half of the FDA-approved anti-AIDS drugs target RT; these comprise eight

Additional Supporting Information may be found in the online version of this article.

Abbreviations: DCC, dynamic cross correlation; EEMs, excision-enhancing mutations; FEL, free energy landscape; NNRTI, non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; NVP, nevirapine; MD, molecular dynamics; PCA, principal component analysis; RT, reverse transcriptase; SASA, solvent-accessible surface area.

*Correspondence to: K. Das; Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, New Jersey 08854. E-mail: kalyan@cabm.rutgers.edu

Received 19 June 2013; Revised 10 October 2013; Accepted 21 October 2013
Published online 22 November 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24460

nucleoside/nucleotide RT inhibitor (NRTIs) and five non-nucleoside RT inhibitors (NNRTIs). NRTIs are incorporated into the growing viral DNA strand by RT and act as chain terminators, whereas NNRTIs are allosteric inhibitors of HIV-1 RT.

RT undergoes conformational changes to carry out its functions. An NNRTI binds to a hydrophobic pocket adjacent to the polymerase active site and allosterically inhibits DNA polymerization. Nevirapine (NVP), a dipyridodiazepinone derivative, was the first clinically approved NNRTI drug; NVP serves as a prototype for this class of inhibitors. An NNRTI inhibits RT by restricting the conformational adaptability of the enzyme.[1] Therefore, it is important to understand the impact of an NNRTI on the dynamic states of RT. Several molecular dynamics (MD) studies of RT and RT-NNRTI complexes have been carried out in the past. A multi-copy MD simulation study of RT-NVP binary complex[4] favored the "molecular arthritis model"[1] of NNRTI inhibition, that is, the binding of an NNRTI restricts the flexibility of thumb. In contrast, a long-range MD simulation study of an RT-efavirenz binary complex suggested that the thumb subdomain also folds down into the nucleic acid-binding cleft of RT.[5] However, experimental structures have not revealed the existence of such a thumb-down conformation of an RT-NNRTI complex, and such a transient state would not permit the binding of nucleic acid. The flexibility of RT is highly dependent on its liganded state. MD simulations of RT-NNRTI complex and unliganded (apo) RT revealed that apo-RT samples a larger conformational space, which is restricted upon the binding of an NNRTI.[6] A recent 2.3 μs time scale MD simulation of apo RT revealed an extended conformation of the fingers and thumb domain,[7] which was not evident earlier. Further the NNRTI pocket, which is inexistent in apo RT structures, did not open up even at such long timescale. This finding favors an "induced fit" mechanism-driven NNRTI binding at the time scale observed. A Gaussian network model based coarse-grained dynamics study of RT in the presence and absence of an NNRTI suggested that RT undergoes large-scale cooperative motion, and the binding of an NNRTI induces suppression of mobility of the p66 thumb.[8] Later, an anisotropic network model (ANM)-based elastic network modeling study proposed that changes in the direction of domain movements rather than suppression of domain motions are responsible for inhibition by an NNRTI.[9] Another study combining ANM and clustering of X-ray crystal structures of HIV-1 RT proposed that hydrophobic core mutations rescue the active-state dynamics of RT.[10]

Experimental biophysical studies have also probed the effects of NNRTI binding on the dynamic states of RT and RT-nucleic acid complexes. A hydrogen-deuterium exchange coupled with mass spectrometry (HDX MS) study of apo RT[11] and RT-efavirenz binary complex revealed that binding of efavirenz reduces the flexibility of multiple regions of RT, and that allosteric impacts of the NNRTI binding are propagated to the distant RNase H domain and the p51 subunit.[12] A single molecule FRET study of RT-DNA (or RT-RNA/DNA) complex showed that (i) RT constantly flips and slides over a double-stranded nucleic acid, (ii) binding of a dNTP stabilizes RT-nucleic acid complex in a polymerase-competent mode, and (iii) binding of an NNRTI enhances the switching frequency of RT, presumably by loosening the fingers–thumb clamp that grasps the nucleic acid.[13,14] Functions of RT at the molecular level are often accompanied with changes in conformation;[15] therefore, it is important to assess the impacts of NNRTI binding on the dynamic states of an RT-nucleic acid complex.

Resistance to existing drugs is a critical problem in the treatment of HIV-1 infections. Primary NRTI-resistance mutations either facilitate discrimination of drugs from dNTP substrate (exclusion) or help removal of an incorporated NRTI from the DNA primer (excision).[16] The set of mutations that enhance excision are called thymidine analog mutations or excision-enhancing mutations (EEMs).[17] Mutations in and around the NNRTI pocket emerge to confer resistance to NNRTIs by either reducing the binding or altering the dynamics of entry/exit of a drug.[18] Earlier clinical isolates were generally sequenced only for the fingers and palm subdomains. Relatively recent sequencing of complete RT from clinical isolates revealed that mutations in the spatially distant connection subdomain and RNase H domain also contribute to NRTI and NNRTI resistance.[19–21] The connection mutations N348I or T369I in combination with primary EEMs or NNRTI-resistance mutations enhance NRTI or NNRTI resistance, respectively[21,22]. Also, the N348I/T369I combination reduces the susceptibility of NNRTIs, and reduces replication capacity of RT.[22] The double mutant confers as high as 60-fold resistance to NVP. Although several biochemical and kinetic studies have focused on understanding the mechanism by which the connection mutations confer resistance to NNRTIs,[23–25] the structural roles of these mutations in imparting resistance largely remains elusive. Therefore, understanding of (i) the overall structural consequence of the connection mutations N348I and T369I on RT, and (ii) how these spatially distant mutations trigger NNRTI and NRTI resistance have important functional implications.

Here, we report an MD study of RT-DNA, RT-DNA-NVP[26], and N348I/T369I mutant RT-DNA-NVP complexes. The study was carried out (i) to understand the conformational dynamics of RT-DNA complexes, (ii) to elucidate if the RT-DNA-NVP complex would access a catalytically competent conformational state during the course of the simulation, and (iii) to explore plausible long-range impacts of the connection subdomain

mutations in enhancing drug resistance. To our knowledge, this is the first MD simulation study of an RT-DNA-NNRTI ternary complex or of an RT complex carrying clinically relevant connection subdomain mutation. MD simulations coupled with dynamic protein structure network (PSN) analyses provide a molecular level understanding of allosteric networks in RT.

MATERIALS AND METHODS

Initial coordinates and MD system setup

The crystal structure of RT-DNA-NVP (PDB: 3V81)[26] that represents a polymerase incompetent state served as a ternary complex for our simulation study, and the clinically relevant connection subdomain mutations N348I and T369I were introduced computationally into the crystal structure of RT-DNA-NVP complex to generate the starting model of the N348I/T369I RT-DNA-NVP complex. The crystal structure of RT-DNA-dTTP (PDB: 1RTD)[27] that represents a polymerase competent state was used to construct a binary complex of RT-DNA by removing the incoming dTTP substrate. This catalytic RT-DNA state is structurally and functionally different from the polymerase-incompetent state of RT-DNA in the presence of an NNRTI.

Swiss PDBV tool[28] was used for modeling the mutations, by optimizing the rotamer based on the lowest clash score. The coordinates for the non-resolved (p51 loop 218–230) region were assigned manually by copying coordinates from the crystal structure of RT-DNA-RNA (PDB: 1HYS)[29] after appropriate structural superposition. The mutated residue Q258C and the non-standard nucleotide bases engineered for cross-linking were reverted to standard residue and nucleotide, respectively. The crystallographic waters were removed, and hydrogens were added using pdb2gmx tool of Gromacs[30] assuming a dominant protonation state for the titratable residues. The crystallographic coordinates of the Mg^{2+} ions present in RT-DNA complex (1RTD) were retained. However, crystal structure of RT-DNA-NVP and N348I/T369I RT-DNA-NVP did not contain active site metal ions (Mg^{2+}), pertaining to the experimental observations that no metal ion binds to the catalytic aspartate when RT is complexed with an NNRTI[31].

NVP was parameterized according to generalized amber force field (GAFF) parameters[32] using ACPYPE,[33] a program that uses a Python parser as an interface for the Antechamber program. Single point AM1-BCC partial charges[34] for the bound NVP were computed using Chimera.[35] All-atom MD simulations were performed in GROMACS (version 4.5.4)[30] using the Amber 03 force field.[36] The system was solvated using the TIP3P water model[37] in a dodecahedral cell keeping a minimum distance of 9 Å between the solute and each face of the box. The total charge was neutralized by the addition of counter ions. An ionic concentration of 0.15 M NaCl, which mimics a physiological concentration, was maintained. Leap-frog algorithm was employed for integrating Newton's equation of motion. A modified Berendsen thermostat[38] and a Parrinello–Rahman barostat[39] were employed for temperature and pressure coupling, respectively. Long-range electrostatic interactions were calculated using particle mesh Ewald[40] method, and the non-bonded interactions were truncated at 12 Å cutoff. The time step for integration was 2 fs, and the coordinates of all atoms were written out every 2 ps for analysis. Holonomic bond constraints were employed using the LINCS algorithm.[41]

Prior to simulation, the system was energy minimized using 3000 steps of steepest descent. A 200-ps NVT equilibration followed by 200 ps of NPT equilibration (300 K, 1 bar; Berendsen coupling scheme), with initial velocities assigned randomly from a Maxwell–Boltzmann distribution, was carried out to allow the solvent to equilibrate around the solute. A 50-ns NPT production simulation run was initiated from the final snapshot of the equilibration run.

Essential dynamics (ED) by PCA

The essential subspace representing the most correlated dominant motions that are fundamental to the activity of the protein was sampled using essential dynamics. Essential dynamics[42] uses a multivariate statistical approach termed principal component analysis (PCA) to extract 'essential' motions from a trajectory fitted to a reference structure with all the translational and rotational motions removed. The internal motion in the trajectory is mathematically represented as $X(t)$, where X represents the $3N$-dimensional column vector of the atomic coordinates. A variance–covariance matrix of positional fluctuations is constructed

$$C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \quad (1)$$

where $x_1, x_2, x_3, \ldots, x_n$ are the mass-weighted Cartesian coordinates of an N particle system and $\langle\ \rangle$ denotes an ensemble average over the time frames, and C represents the symmetric matrix. The symmetric 3N × 3N matrix (C) thus obtained is diagonalized by an orthogonal coordinate transformation matrix (R).

$$R^{\mathrm{T}} C R = \mathrm{diag}(\lambda_{1,}\lambda_{2,} \ldots \lambda_{n,}), \text{where } \lambda_1 \geq \lambda_2 \geq \lambda_2 \cdots \geq \lambda_{3n,} \quad (2)$$

λ signifies the diagonal elements (eigenvalues) of the matrix and the columns of R represents the eigenvectors (principal modes). The eigenvalues λ correspond to the mean-square fluctuation, and describes the magnitude of motion, while eigenvector describes the direction of motion and R^{T} is the transpose of R. In general, the eigenvectors are sorted based on decreasing order of

eigenvalues, and the top principal modes (eigenvectors) account for the collective global motions in the system.

The motion described by a principal mode can be visualized by projecting the trajectory on the principal modes to yield the principal coordinates.

$$q = R^{T}(X(t) - \langle X \rangle) \quad (3)$$

Essential dynamics was carried out using the backbone atoms, to overcome apparent artificial correlations resulting from fast side-chain movements. The extreme motions obtained by projecting the first eigenvector onto each trajectory served as input for the program DynDom,[43] which identifies dynamic domains, hinge axes, and hinge-bending residues.

Characterization of energy landscape

A protein in solution exists as a dynamic ensemble of inter-converting conformations. Conformational analysis provides insights for structural transition and helps identify distinct conformational basins that help in establishing structure–function relationship. Such landscapes could unravel meta-stable states that have not been observed experimentally. However, computing free energy landscapes (FELs) as a function of 3N-6 coordinates is computationally prohibitive. Hence, a low-dimensional FEL calculation was carried out using the first two principal components obtained from individual trajectories. The first two PCs of the respective systems served as reaction coordinates to generate a two-dimensional FEL plot. This was implemented using the g_sham module of Gromacs.[30]

$$G\alpha = -kT \ln \frac{P(q_\alpha)}{P_{max}(q)} \quad (4)$$

where k is the Boltzmann constant, T is the temperature of simulation, $P(q_\alpha)$ is an estimate of the probability density function obtained from a histogram of the MD data, and $P_{max}(q)$ is the probability of the most populated state. Using the projections of the trajectory along the first and second principal components (PC1, PC2) as reaction coordinates (q_i and q_j), a joint probability distribution $P(q_i,q_j)$ of the system was obtained.

Dynamic cross correlation maps

Cross-correlated fluctuations between any two pair of C_α atoms were calculated by the Bio3D package[44] using the equation

$$C_{ij} = \frac{\langle \Delta r_i . \Delta r_j \rangle}{\sqrt{\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle}} \quad (5)$$

where i and j correspond to any two C_α atoms, displacement vectors Δr_i and Δr_j represents the displacement of the ith atom and jth atom from their mean positions, and the symbol $\langle\ \rangle$ represents the ensemble average over the entire trajectory. The magnitudes of all pairwise cross-correlation coefficients are obtained as a matrix and are graphically represented as a dynamic cross-correlation (DCC) map. The fluctuations are designated as completely correlated (same period and phase) if $C(i, j) = 1$, completely anti-correlated (same period but opposite phase) if $C(i, j) = -1$.

PSN analysis

The network paradigm perceives each amino acid in the protein structure as a node and the strength of the non-covalent interactions between the node pair as links (edges). The weights (interaction strength) for the edges were computed for all frames constituting the MD trajectories using the equation

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} \times 100 \quad (6)$$

where I_{ij} is the strength of the interaction between residue i and j, n_{ij} is the number of distinct atom pairs between residues i and j coming within a distance of 4.5 Å, and N_i and N_j are the normalization factors (NF) for residues i and j, which take into account the differences in size of the different nodes and their propensity to make the maximum number of contacts with other nodes in protein structures.[45] Each node pair with an I_{ij} value greater than or equal to a given I_{min} cutoff is connected by an edge. At high I_{min} cutoffs, only nodes with high numbers of interacting atom pairs will be connected by edges, an indication of stronger inter-residue interactions. Multiple protein structural networks were constructed by varying I_{min} cutoffs (3, 3.5, and 4 Å) to identify hubs (nodes having high degree of connectivity). The percentage of interaction in a hub is quantified using the formula

$$I_i = \frac{n_{ij}}{N_i} \times 100 \quad (7)$$

where I_i is the hub interaction percentage of node i, n_{ij} is the number of side-chain atom pairs within a given distance cutoff, and $N_i$is the NF of residue i. The networks thus obtained at various I_{min} cutoffs were analyzed for cliques, communities, hubs, and communication pathways. The shortest non-covalently connected path(s) between a selected pair of residues in a trajectory was calculated using Dijkstra's algorithm.[46] Identification of the shortest path involves a search for all possible shortest paths between the selected residues from the PSN followed by identification of an optimal path that contains at least one dynamically correlated residue with the selected pair. PSN analysis was carried out using

Wordom,[47] hubs in the network were visualized using xPyder,[48] cliques and communities in networks were identified using CFinder,[49] and other associated network parameters were computed using Cytoscape.[50] It should be noted that nucleic acid was not considered for network analysis, because the underlying nucleic acid recognition mechanism in HIV-1 RT is an indirect readout mechanism, which is dependent on its shape rather than sequence. A recent kinetic study in the absence of a nucleic acid has experimentally demonstrated a reduced binding of NVP to RT by N348I mutation.[25] Accordingly, it was presumed that allosteric signaling path in RT is predominantly independent of the nucleic acid.

RESULTS AND DISCUSSION

MD simulations of RT-DNA, RT-DNA-NVP, and N348I/T369I RT-DNA-NVP complexes

We carried out 50-ns simulations of RT-DNA (PDB: 1RTD),[27] RT-DNA-NVP (PDB: ID 3V81),[26] and the connection subdomain double mutant N348I/T369I RT-DNA-NVP structures, under conditions described in the experimental section. The starting model of the N348I/T369I complex was obtained by introducing the mutations in both the p66 and p51 subunits of RT-DNA-NVP structure. Comparative analysis of the trajectories of RT-DNA versus RT-DNA-NVP and RT-DNA-NVP versus N348I/T369I RT-DNA-NVP were carried out for understanding how NVP binding alters the dynamics of RT and how the distal mutations in the connection subdomain contribute to NVP resistance, respectively. The RT-DNA cross-link was removed to allow unrestrained motion during the course of simulation. The structural stability of each MD trajectory was ascertained by analyzing RMSD (Supporting Information Fig. S1A) and radius of gyration (R_g, an indicator of protein compactness) plots (Supporting Information Fig. S1B). The RMSD of the protein backbone with respect to the corresponding starting X-ray structure revealed an initial spike of ~3 Å within the first 5 ns of MD simulation, after which all the three systems remained stable throughout the simulation. Hence, 5–50 ns trajectories were adopted for the production stages. R_g remained largely constant throughout the course of the simulation suggesting no loss in structural compactness or structural meltdown (Supporting Information Fig. S1B).

Structural rearrangements in HIV-1 RT were analyzed on pico to nano second time scales. The p51 subunit was relatively rigid during the course of MD simulation for all three complexes, whereas the p66 subunit exhibited significant flexibility (Supporting Information Fig. S1C). Therefore, all of the structural elements discussed hereafter are in p66 unless otherwise stated. The root-mean-square fluctuations computed from MD simulations were translated to simulated B-factors using a transformation

$$B = \frac{8\pi^2}{3} \times (\mathrm{RMSF})^2 \quad (8)$$

A high degree of correlation in their relative magnitudes when compared with crystallographic B-factor, indicates that the MD simulation recapitulates the inherent flexibility in RT (Supporting Information Fig. S2). The mean standard error for the computed B-factors were determined using a block averaging approach.

Dynamic cross-correlated motion between C_α atom pairs was used to identify regions of RT that undergo concerted motions. Residue pairs with correlations greater than ± 0.5 were considered to be significant and are highlighted in the DCC maps provided in (Supporting Information Fig. S3). DCC analysis revealed subtle changes in correlated and anti-correlated motions among the subdomains and domains of RT upon NVP binding. A correlated motion between the two segments of the fingers subdomain was evident in both RT-DNA and RT-DNA-NVP complexes. This correlation may represent the dynamic motion of the fingers that opens and closes to bind a dNTP.[27] The extent of correlation/anti-correlation is greatly enhanced in N348I/T369I RT-DNA-NVP complex, indicating that these connection mutations have pronounced effects on the dynamic states of RT. A weak anti-correlated motion between fingers and thumb subdomains that appeared upon NVP binding was further enhanced in the N348I/T369I RT-DNA-NVP complex (Supporting Information Figs. S3B, S3C). Additionally, the N348I/T369I RT-DNA-NVP complex displayed an anti-correlated motion between fingers and RNase H, which implies that N348I/T369I mutations propagate changes in structure and mobility well beyond its immediate vicinity. Such long-range structural impacts may help in assessing the role of the connection mutations in developing resistance to both NRTIs and NNRTIs.[19–21] A detailed study of long-range allosteric perturbations and communication pathways in RT is discussed later.

NNRTI binding and structural rearrangements of the polymerase domain

Structural rearrangements and conformational transitions orchestrate the biochemical functions of RT.[11,14–16] Therefore, an analysis of the structural rearrangements within the polymerase domain was carried out in the course of MD simulations. The relative separation between the thumb and fingers subdomains defines the opening and closing of the clamp that holds and positions a dsDNA or RNA/DNA template-primer at the polymerase active site. The precise positioning of the primer 3′ -end with respect to the polymerase active site (D110, D185, and D186) is required for the catalytic

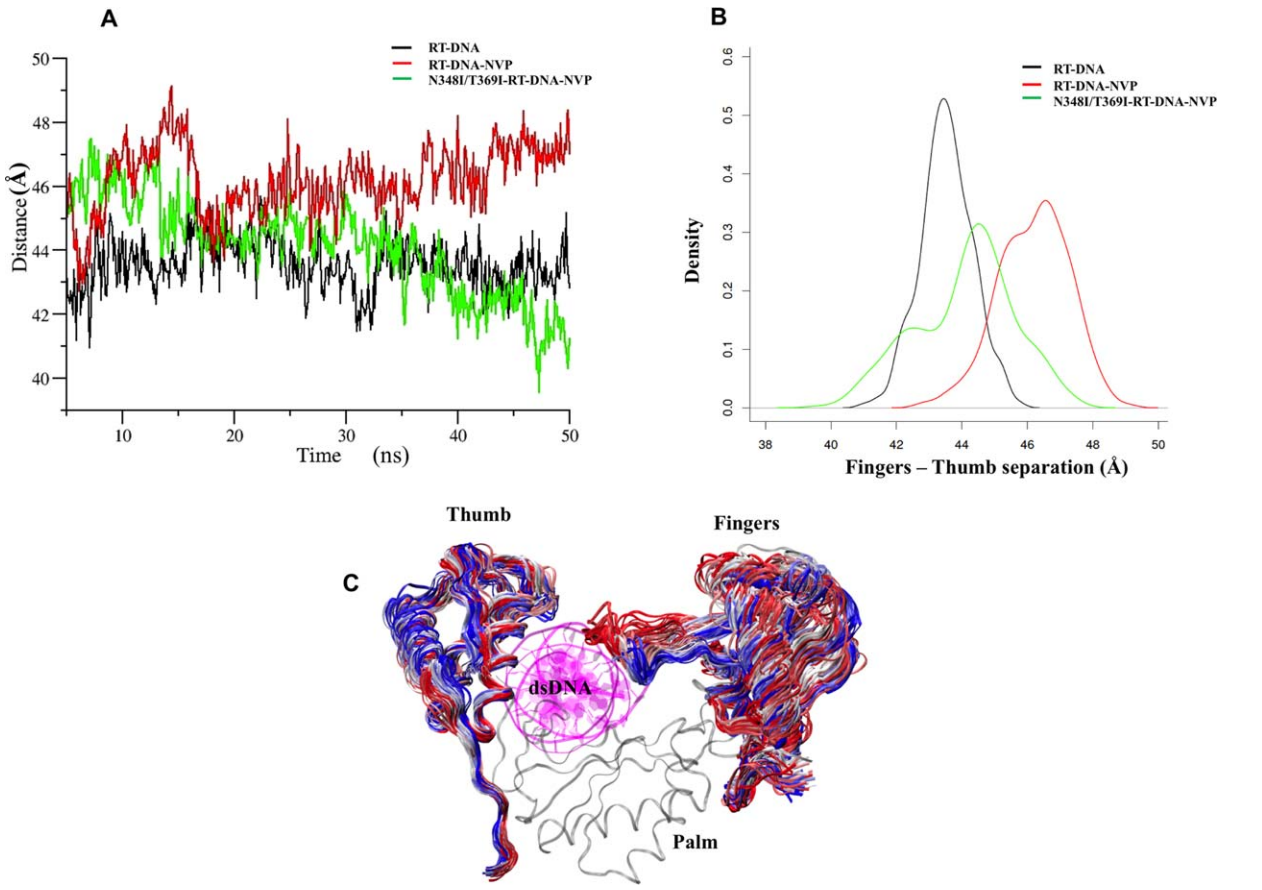

Figure 1

Opening and closing of the cleft between the fingers and thumb subdomains during the course of MD simulations. (**A**) The distance between the centers of mass (COM) of fingers and COM of thumb subdomains in the course of 50-ns MD simulations of RT-DNA (black), RT-DNA-NVP (red), and N348I/T369I RT-DNA-NVP (green) structures. (B) Distribution profile based on fingers–thumb separation distance. (**C**) Ensembles of fingers and thumb positions obtained from the MD simulation of the RT-DNA-NVP complex show the flexibility of the thumb and finger subdomains. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

addition of a nucleotide. To assess the structural impacts of NVP binding on RT, we computed the relative distances between the (i) centers of mass of fingers and thumb and (ii) the centers of mass of the polymerase active site residues and the primer 3′ -end nucleotide for the entire length of the simulations. The separation between thumb and fingers fluctuates about a mean distance of 43 Å in the polymerase-competent RT-DNA complex suggesting a stable interaction between RT and nucleic acid [Fig. 1(A)]. The fingers–thumb separation in RT-DNA-NVP complex reached a maximum around ∼10–15 ns timescale, and this state mimics the thumb hyper-extended conformation observed in the X-ray structures of RT-NNRTI binary complexes. A probability density distribution plot [Fig. 1(B)] of fingers–thumb separation revealed a unimodal distribution with a peak value at ∼43 Å for RT-DNA, and a peak at ∼47 Å for RT-DNA-NVP complex and a bimodal distribution with two peaks at ∼42 and ∼45 Å for N348I/T369I RT-DNA-NVP complex. Superposition of the MD ensemble structures of RT-DNA-NVP complex [Fig. 1(C)] demonstrates that the cleft could adapt conformational states, as seen in RT-DNA complex.

Opening of the cleft essentially results in loss of nonbonded interactions between nucleic acid and RT. The average interaction energy, $E_{\text{interaction}}$ (E_{vdw} + E_{ele}), between RT and dsDNA was computed for the RT-DNA and RT-DNA-NVP complexes; $E_{\text{interaction}}$ of −452 ± 29.82 kcal/mol for the RT-DNA-NVP and −528 ± 45.24 kcal/mol for the RT-DNA complex, respectively, indicates that the interaction between RT and DNA is decreased upon NVP binding. This loss of interaction is consistent with the single molecule FRET study that demonstrated loosening of the fingers–thumb clamp and enhanced dissociation of nucleic acid from RT upon NVP binding[13].

Crystal structure of the RT-DNA-NVP complex showed that the binding of NVP repositions the primer 3′ -end away from the polymerase active site.[26] In order

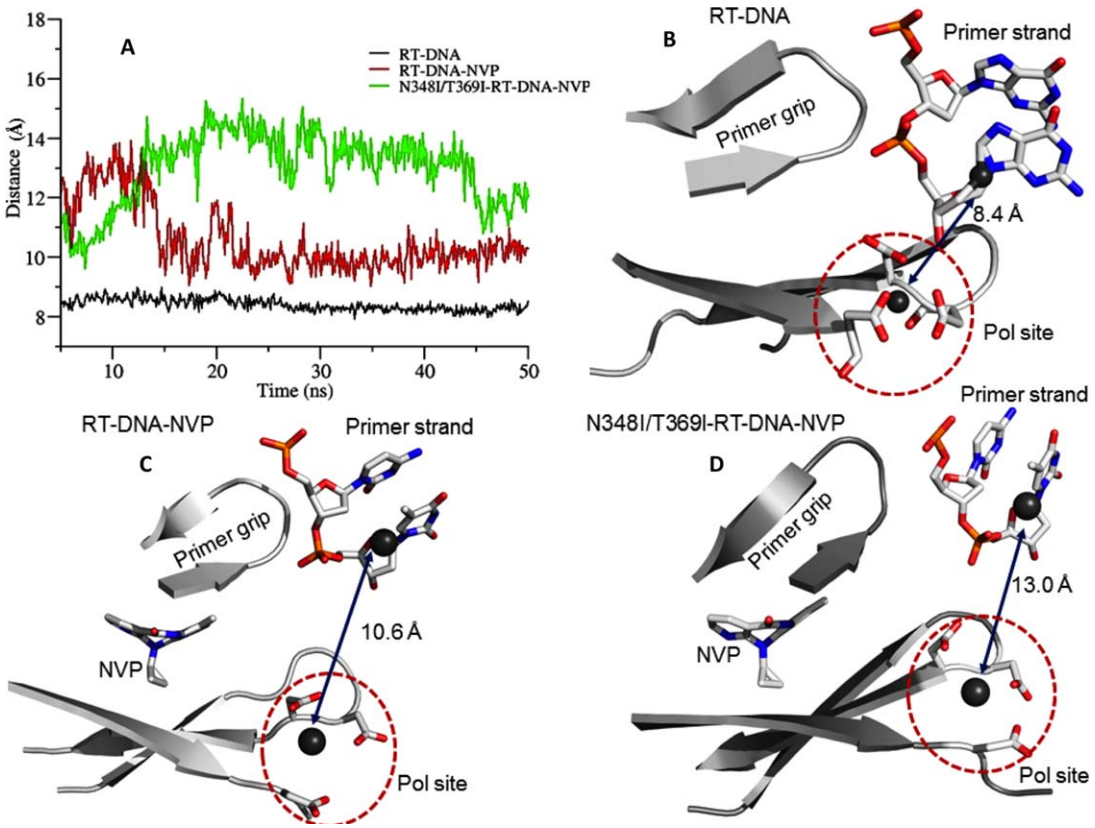

Figure 2

Probing the catalytically-competent states of RT as a function of its distance between the polymerase active site and the primer 3′-terminal nucleotide during the course of MD simulations. (**A**) The distance between the COM of the primer 3′-end nucleotide and the COM of the polymerase active site. (**B**) Polymerase active site of the RT-DNA complex is in a catalytically competent mode. In the RT-DNA-NVP complex (**C**) or N348I/T369I mutant RT-DNA-NVP complex (**D**), the distance between the polymerase active site and the primer terminal nucleotide is increased. The distances listed in figure panels B, C, and D are the average distance calculated from panel A. At any instance of the simulation, the nevirapine-bound complexes the distance did not attain a value comparable to that observed in a catalytically-competent structure. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

to find if the 3′-end can assume a catalytically-competent conformation, we analyzed the relative positioning of the primer terminal nucleotide with respect to the polymerase active site in the course of MD simulation [Fig. 2(A)]. The distance remained relatively constant at ∼8.4 Å during the MD simulation of RT-DNA, representing a state compatible for dNTP binding and nucleotide incorporation. Simulations of RT-DNA-NVP and N348I/T369I RT-DNA-NVP complexes revealed that in NVP-bound structures, the primer 3′-end did not reach the active site at any instant of simulation (Fig. 2). The inability of the primer terminus to attain a polymerase-competent state during the MD simulations of RT-DNA-NVP complexes suggests that the conformational restriction imposed upon NNRTI binding is a primary mechanism of inhibition. Additionally, the side chains of the catalytic aspartates (D110, D185, and D186) sampled distinct sets of dihedral conformations in the simulations of NVP-bound versus unbound structures, which may correlate with the fact that no metal chelation is observed at the polymerase active site in any of the reported structure of RT-NNRTI complexes.[31] A dNTP binds to the RT-DNA complex at the active site (N site) by (i) base stacking with the nucleic acid, (ii) base pairing with the template overhang, and (iii) metal chelation of the triphosphate group. The inability of the nucleic acid to reach the active site and the inability of the active site to chelate metal ions suggests that a dNTP would not be able to bind at the N site when RT is bound to an NNRTI. A recent kinetic study using an incremental isothermal titration calorimetry technique revealed that dNTPs in fact do not bind to RT-DNA-NNRTI complexes.[51] Overall, analysis of our MD simulations revealed that binding of an NNRTI has multiple effects, primarily caused by repositioning of the primer grip and distortion of the active site triad; however, the fingers subdomain remained flexible in the RT-DNA-NVP complexes (Supporting Information Fig. S1C).

Essential dynamics analysis and global motions of RT

PCA-based essential dynamics[42] was used for extracting large-scale concerted motions from the MD trajectories. The cumulative fraction of variance explained by the first 10 PCs are shown in Supporting Information Fig. S4. The first 10 PCs accounted for 24.7, 28.8, and 44.2% of the total variance in conformational dynamics seen in RT-DNA, RT-DNA-NVP, and N348I/T369I RT-DNA-NVP complexes, respectively. Comparison of the essential subspace spanned by RT-DNA versus RT-DNA-NVP trajectory and RT-DNA-NVP versus N348I/T369I RT-DNA-NVP trajectory were carried out by cross projecting one set of eigenvectors onto another in a 2D plane (Supporting Information Fig. S5). The cross projections shows that the essential eigenvectors obtained from RT-DNA and RT-DNA-NVP trajectories have some overlap (Supporting Information Fig. S5), indicating that the RT-DNA-NVP complex samples a part of the essential space seen in RT-DNA complex. Similarly, cross projection of the RT-DNA-NVP trajectory over the mutant N348I/T369I RT-DNA-NVP trajectory reveals a larger spread, indicative of an enhancement in sampling space of the mutant form.

Cosine content analysis[52] was carried out using g_analyze tool of gromacs to check for convergence using the top 10 eigenvectors and their values are reported in **SI** Table 1. A high cosine content value was obtained for RT-DNA and RT-DNA-NVP complexes spanning the entire 5–50 ns time; however, cosine content values calculated for 5–25 ns and 30–50 ns reveled low values, which lead us to conclude that the high cosine content obtained from 5 to 50 ns is not indicative of random diffusion, but rather due a conformational transition to another metastable state. The RMSD plot (Supporting Information Fig. S1A) also corroborates a structural transition occurring between ~31 and 33 ns time scale for RT-DNA complex and around ~12–15 ns time scale for RT-DNA-NVP complex. It should also be noted that a low cosine does not ensure convergence.[52] Arguably, our MD trajectories indicate that sampling has been adequate to validate the dynamics at the observed timescale, but does not guarantee convergence. In general, quantitative estimation of convergence requires longer simulation time, but how long is clearly unpredictable.[53] For a complex and flexible system like RT, convergence may not even be conceivable at currently attainable timescales (μs second simulation of RT using Anton[7]).

The extreme conformations obtained from principal component 1 (PC1) were used to perform a rigid body analysis using DynDom[43] that identifies dynamic domains, hinge axes, and hinge-bending regions. DynDom analysis of the global motion derived from PC1 for RT-DNA revealed an axis associated with the closing and opening of fingers (**SI Movie 1**). The closing and opening of the fingers is essential for binding of dNTP, release of pyrophosphate, and translocation of nucleic acid. For example, a single molecule fluorescence study of T7 DNA polymerase revealed that the fingers fluctuation is associated with the binding of a correctly base-paired dNTP.[54] Like other DNA polymerases, HIV-1 RT also requires the closing and opening of the fingers subdomain.[27] Our MD simulations seemingly captured the swivel motion of closing and opening of fingers in all three structures. However, the directions and amplitudes of the open/close motions of the fingers were different for each of the three complexes (**Movies S1, S2, S3**). The fingers subdomain swiveled approximately 21° and 25° about their respective axes in RT-DNA and RT-DNA-NVP complexes, whereas this motion was restricted to ~5° in the N348I/T369I RT-DNA-NVP complex.

A second hinge axis (axis 2) was identified for the RT-DNA-NVP and N348I/T369I RT-DNA-NVP complexes (**Movies S2, S3**), and the axis passed through the NNRTI-binding site. This hinge axis represented an *en bloc* bending motion of a large segment of RT including connection and RNase H. This hinge-bending motion is responsible for repositioning of the RNase H domain with respect to the polymerase active site; a maximum of ~14° and 10° rotation about axis 2 was observed for RT-DNA-NVP and N348I/T369I RT-DNA-NVP complexes, respectively; no significant rearrangement of the RNase H domain was evident in the course of simulation of the RT-DNA complex. The motion about axis 2 may account for the altered RNase H cleavage specificity upon NVP binding[55].

DynDom analysis also predicted a third hinge axis for the N348I/T369I RT-DNA-NVP complex only. This axis traversed through the p51 subunit, and the hinge motion caused p51 subunit to swivel by ~10° and repositioned the RNase H domain (**Movie S3**). Although the N348I/T369I mutations were present in both subunits, axis 3 passes through a region adjacent to the mutation sites in p51. Therefore, we speculate that the mutations in p51 might have a larger impact on RNase H activity. This is in agreement with a subunit-specific kinetic study of N348I-mutant RT showing that the mutation in the p51 subunit decreases RNase H activity.[25] A recent study hypothesized that the loss of a hydrogen bond between N348 and Y427 in p51 by N348I mutation may be responsible for the decreased RNase H activity by the mutant RT.[56] The effects of the N348I/T369I mutations in both subunits are discussed in "Network analysis" section.

Characterization of the energy landscape

Apart from understanding the structural dynamics of HIV-1 RT, it is equally important to elucidate the energy landscape of RT in the three complexes. In general, internal motion of a protein in solution is driven

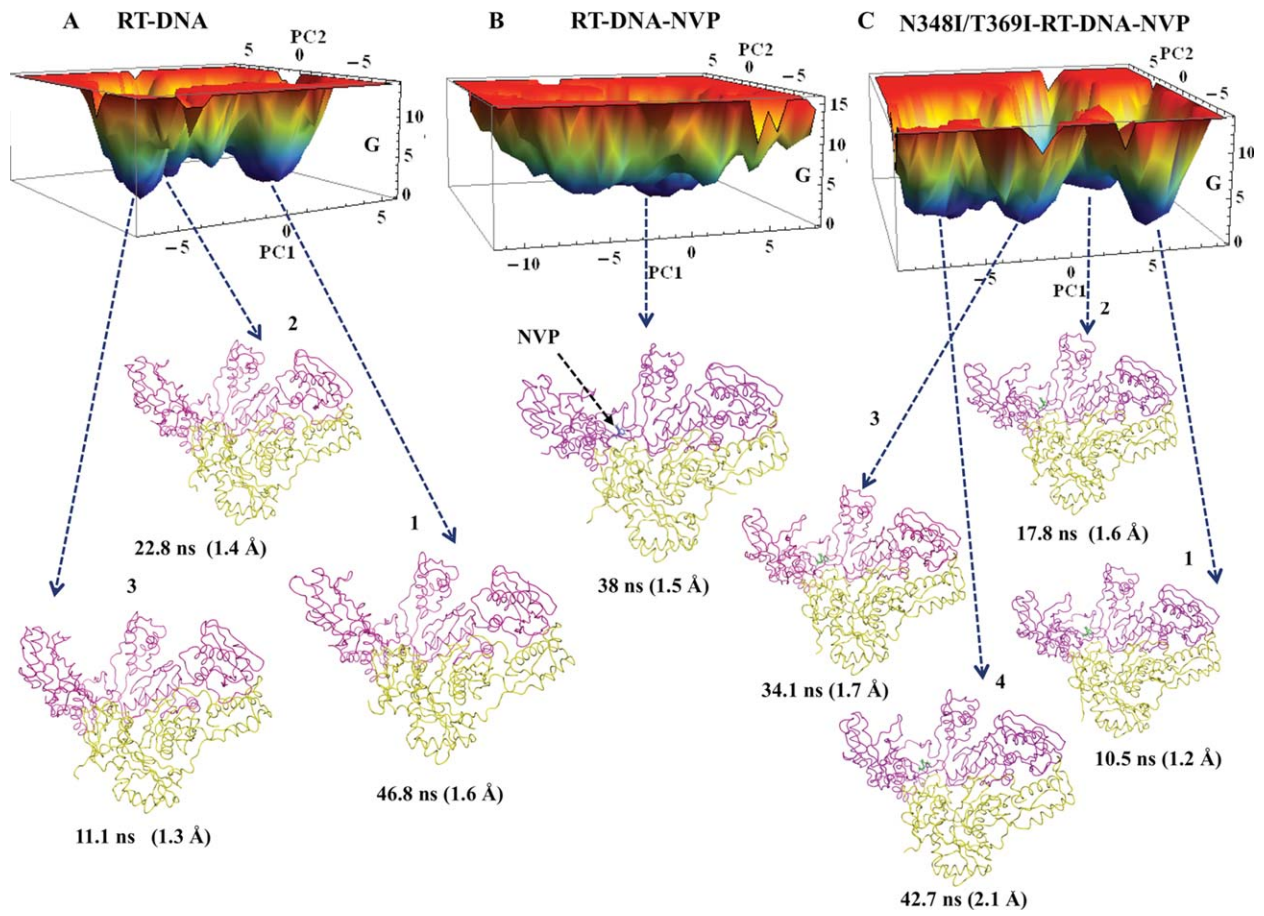

Figure 3

Free energy landscape obtained by projecting the individual trajectories on to their respective principal components (PC1 and PC2). Free energy landscape of RT-DNA (**A**), RT-DNA-NVP (**B**), and N348I/T369I mutant RT-DNA-NVP (**C**) complexes. Representative conformations of RT from each major low energy basin are plotted below. The conformational states of RT corresponding to individual basins are shown; the time scale and rmsd from the starting structures are listed.

thermodynamically, which allows sampling of metastable states in the process of overcoming an energy barrier. Hence, a conformational space based low-dimensional FEL of individual system was generated using PC1 and PC2 as reaction coordinates[57] (Fig. 3). It should be noted that the low-dimensional FEL representations might not be precise enough for identifying free-energy barriers and metastable states accurately;[58] however, these representations can be extremely useful to describe the energy landscape and the conformational characteristics of RT in different basins.

The energy landscape for each complex differ significantly from the other two, which is indicative of conformational heterogeneity exhibited by RT in the three different complexes. The FEL of RT-DNA revealed a rugged surface[59] with three distinct minima [Fig. 3(A)]. On the contrary, the RT-DNA-NVP complex was trapped in a single energy basin with a broad and relatively smooth FEL [Fig. 3(B)]. Such a reduction in the accessible conformational freedom of RT might generate a drastic decrease in entropy. This finding is in agreement with a recent HDX study that revealed large-scale rigidification of RT in the presence of efavirenz.[12] The HDX study of apo RT[11] showed that the palm, thumb, connection, and the RNase H domain were significantly more flexible in solution than that expected from crystal structures. This enhanced flexibility of RT in solution may account for the multiple minima evinced in the FEL of RT-DNA complex [Fig. 3(A)]. Rigidification of RT upon binding of an NNRTI appears to be responsible for the conformational biasing of RT to a highly populated single energy state as revealed from the FEL plot [Fig. 3(B)]. A shift in the ensemble population of the RT-DNA complex to a single energy state upon binding of NVP reaffirms the conformational bias imposed by NNRTI binding, which forbids the RT-DNA complex from sampling functionally relevant conformational states.[12] On a related note, MD studies on HCV RNA polymerase have also identified attenuation of motions upon inhibitor binding.[60] Congruence in global motions

hints towards a conserved mechanism among viral polymerases. Similar population shifts in energy landscape upon ligand binding have also been observed for other proteins.[61,62]

The FEL analysis of the N348I/T369I RT-DNA-NVP complex [Fig. 3(C)] was surprisingly distinct from that of RT-DNA-NVP, which suggested that the connection mutations may alter the conformational dynamics of the RT-DNA-NVP complex. An analysis of the NNRTI pocket volume using MDpocket software[63] over the course of the MD simulations revealed no significant changes in the pocket volume of the mutant versus wild-type complex (Supporting Information Fig. S6A). A one-dimensional radial distribution function calculated for the solvent around NVP revealed a similar profile for RT-DNA-NVP and N348I/T369I RT-DNA-NVP complexes (Supporting Information Fig. S6B). The mutants also did not show any significant change in solvent-accessible surface area (ΔSASA) of the NNRTI pocket (Supporting Information Fig. S6C). Absence of any appreciable structural changes at the global and local (NNRTI pocket) level in the N348I/T369I RT-DNA-NVP complex suggested that these mutations might propagate changes by a dynamically driven allosteric mechanism,[64,65] that requires little or no structural change.

PSN analysis and allosteric communication pathways in HIV-1 RT

Allostery by definition involves the binding of an effector molecule at one site that alters the structure and function at a distant site. Two classical phenomenological models, the Monod–Wyman–Changeux (MWC)[66] and Koshland–Némethy–Filmer (KNF)[67], have been used to describe allsotery. However, these models were proposed in the context of structure, and the important role of protein dynamics was overlooked. The role of conformational dynamics in allostery was first proposed by Weber[68] and this led to a new ensemble-based formalism termed as the "conformational selection and population shift" model.[69] NMR experiments have provided direct support to this hypothesis.[61,62,64,70] Given its broader definition, it has been suggested that allostery is an intrinsic property of proteins,[71,72] and occurrence does not require a change in shape.[64,65]

Rearrangements of residue–residue contacts could also contribute to allosteric communications between spatially distal sites in a protein. Therefore, we employed PSN analysis,[73,74] which takes side-chain interactions into account to predict long-range allosteric communications between distal sites. A network representation of protein structures could shed light on (i) assessing long-range impacts of NVP binding on RT-DNA complex and (ii) structural basis for resistance conferred by the distal mutations. Network parameters: hubs (highly connected nodes), cliques (completely connected subgraphs, in which a set of nodes are clustered), and communities (connected cliques) were calculated using the software packages Wordom,[47] xPyder,[48] and CFinder[49]. Nodes (residues) with degree (number of node) >11 were identified as hub (highly connected edges connected to the nodes) residues. The details on the construction of the protein structural network are described in the experimental section.

Hubs identified for all three complexes are displayed in Figure 4(a) and listed in Supporting Information Table S2 and the cliques are shown in Supporting Information Fig. S7. Hub residues are known to be critical in protein folding and stability.[73] Therefore, many of the hub residues are likely to be conserved across homologous proteins, or between different conformational states of a protein. The number of hub residues in p51 is higher than that in the p66 subunit, which reflects (i) higher structural rigidity of p51 and (ii) an asymmetric distribution of hub pattern despite both p51 and p66 having identical amino acid sequence. Comparison of the hub patterns in RT-DNA-NVP versus RT-DNA revealed an increase in the number of hubs upon NVP binding [Fig. 4(A), Supporting Information Table S2]. The hub patterns in the NVP complexes tend to be assortative (hubs preferentially associating with other hubs) whereas, the RT-DNA complex displays a disassortative pattern (Supporting Information Fig. S8). This striking difference in hub pattern is not surprising, because the binding of an NNRTI to RT-DNA essentially enhances rigidity of RT, as observed by HDX.[12] We also found an increase in the number of hubs at the p66/p51 dimer interface for the RT-DNA-NVP complex. An increase in hub patterns, in general, indicates a stronger oligomeric association in protein complexes.[74] The increase in interface hubs upon NVP binding to RT may account for the experimental finding that NNRTI binding enhances p66/p51 dimerization.[75] Interface cluster-forming residues are known to act as hotspots that mediate subunit interaction in oligomeric complexes.[76] We identified an interface hub around residues Trp398 and Trp401 of p66 in the RT-DNA-NVP complex, which is a known hotspot that contains the tryptophan repeat motif (p66 residues 398–414) essential for RT dimerization.[77] Increase in the hub residues between fingers-palm, connection-RNase H, and thumb-connection in the N348I/T369I RT-DNA-NVP complex [Fig. 4(A)] indicated enhanced inter-subdomain communication in the mutant complex.

To understand the allosteric changes in RT upon NVP binding or the effect of connection mutations on the distal sites of RT, we analyzed the differences in the topology of the interaction networks among the three complexes. Higher order network parameters like clique and community patterns were calculated for the entire protein; however, we limit our discussion to the more flexible p66 subunit and the mutated positions in p51. All dynamically stable (if present in >50% of the trajectory frames) cliques and communities for the three

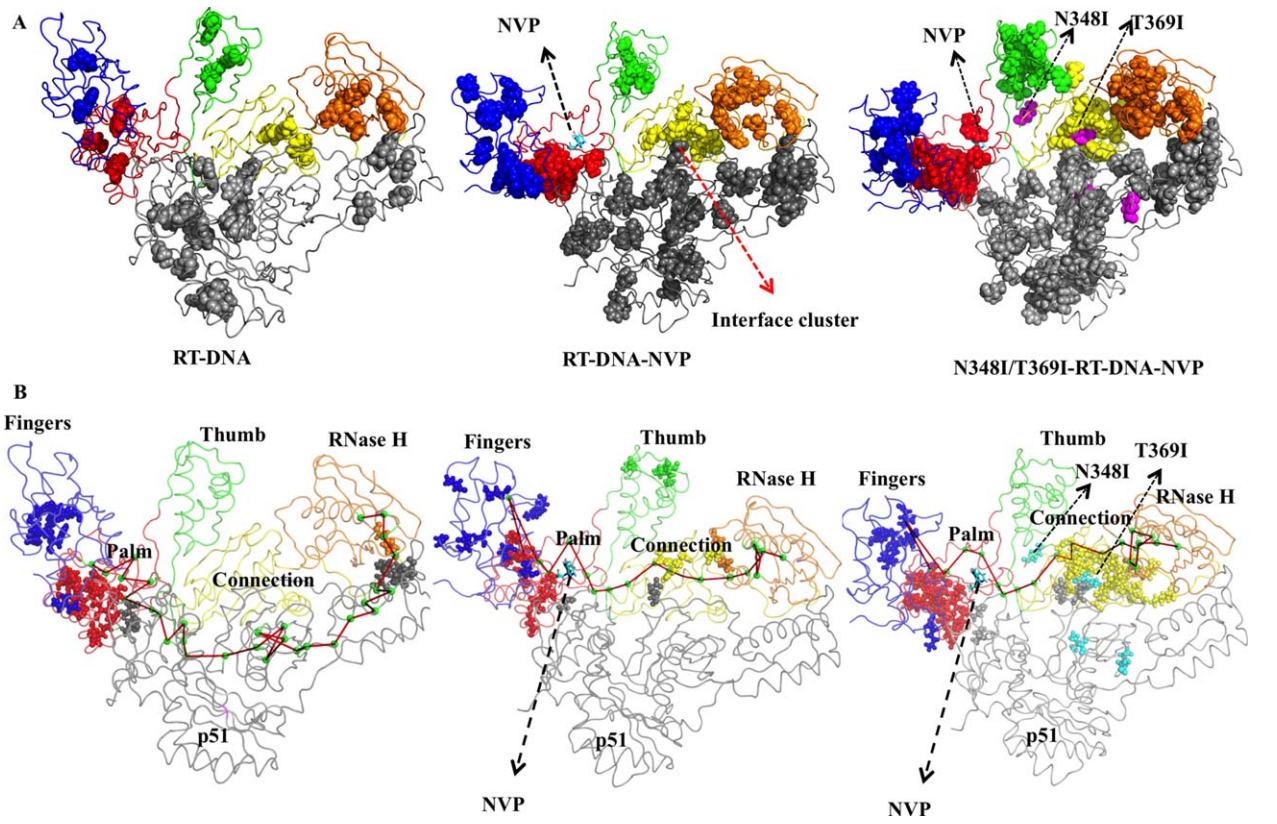

Figure 4

Hubs, communities, and communication paths identified from protein network analysis. (**A**) Hub residues identified for all the three complexes. The hub residues are depicted as van der Waals spheres and color-coded by location (fingers blue; palm red; connection yellow; RNase H orange; p51 gray). (**B**) Calculated communities and the shortest communication paths between the polymerase and RNase H active sites in the three complexes; the communication pathway is represented by red lines. Community-forming residues are shown as van der Waals spheres and color-coded based on the subdomains. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

complexes were used to construct putative communication networks in the three complexes [Fig. 4(B) and Supporting Information Fig. S7]. Binding of NVP caused significant reorganization of the cliques and community patterns both adjacent to the NNRTI pocket and at spatially distant sites [Fig. 4(B)]. Altered community patterns in the fingers and palm subdomains may affect dNTP binding to RT-DNA-NVP complex. New communities appear in the thumb and connection subdomains that may correlate with the observed rigidification of the thumb and the connection upon NNRTI binding.[12]

Polymerization by RT at the Pol site is known to occur in concert with RNase H activity. Studies have established a spatial and temporal interdependence between the two spatially apart active sites.[78] Henceforth, residues D110 (polymerase active site) and D443 (RNase H active site), were used as the end points to identify the shortest communication path. The shortest communication paths between the polymerase site and the RNase H site in all three complexes are listed in Supporting Information Table S3 and shown in Figure 4(B). Our analysis suggests a communication path between the polymerase and RNase H site, which passes through p51 subunit in the RT-DNA complex [Fig. 4(B)]. The communication path traverses via the NNRTI pocket (residues M230, P95, K101, and Y181) suggesting that the NNRTI pocket residues are integral parts of the communication network connecting both active sites of RT. The network predominately passes through the p51 subunit indicating that the p51 subunit, apart from its structural role, may also mediate allosteric communications in RT. Upon NNRTI binding, the p66/p51 dimer becomes more stable, and the communication path passes through the dimer interface. The p51 region N136-E138 that is involved in dimerization and in NNRTI binding is a part of the communication pathway.[79]

Residues N348 and T369 are predominantly surrounded by hydrophobic and aromatic residues and N348I/T369I mutations that enhance hydrophobicity may help form stronger hydrophobic cores surrounding the mutated residues. As a consequence, the N348I/T369I mutant RT-DNA-NVP complex had an increased number of hubs and stronger communities surrounding the

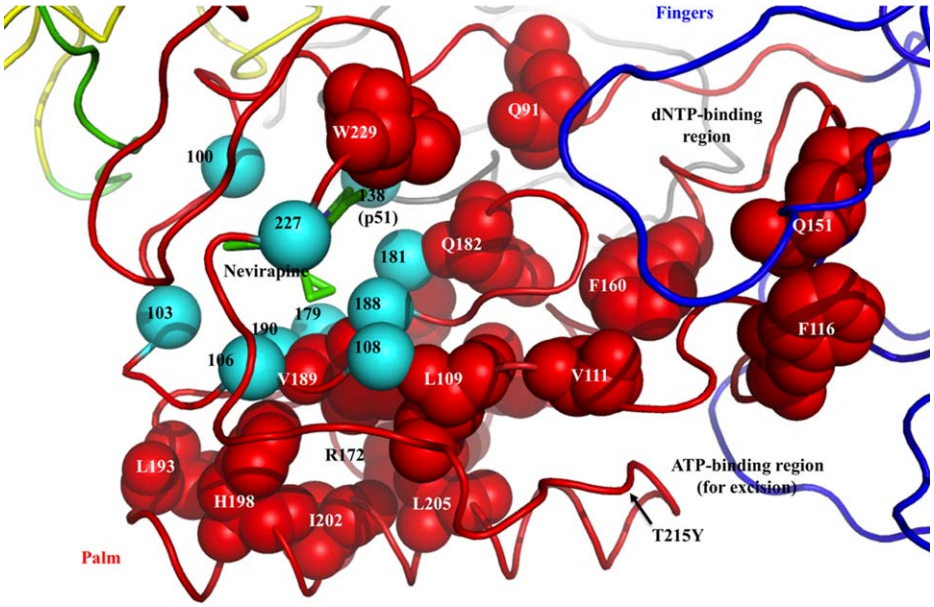

Figure 5

A close view of the distributions of hub residues (red) in the palm subdomain that emerged around the polymerase active site region, ATP-binding site (as an excision agent), and NNRTI pocket in the N348I/T369I mutant RT-DNA-NVP structure. Increases in the number of hub-forming residues imply enhanced structural rigidity around these regions. The common NNRTI-resistance mutation sites are shown as cyan spheres surrounding nevirapine (green). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

mutation sites in both p51 and p66 (Fig. 4). The dense community that is extended from the connection to the RNase H active site reflects enhanced rigidity of the region, which may contribute to the reduced RNase H activity by the connection mutations.[19] The putative communication path between the polymerase and RNase H active sites in the mutant complex is redirected through the p66 connection subdomain.

The connection mutations have been shown to impact DNA polymerization, enhance excision, reduce RNase H activity, and enhance NNRTI resistance.[20–22,24,80] Interestingly, the connection mutations also enhanced the number of hubs and communities in the palm subdomain (Figs. 4 and 5; Supporting Information Table S2), demonstrating enhanced rigidity of the region, which includes the polymerase active site region, ATP-binding site for excision, and NNRTI pocket. A kinetic study revealed that N348I mutation decreases the rate of association (K_{on}) of NVP to RT,[25] which is apparently due to the enhanced rigidity of the NNRTI pocket caused by the connection mutation (Fig. 5). The impact is less on small or flexible NNRTIs[21,22] like efavirenz or etravirine, respectively, that are more resilient to the pocket changes by being more adaptive.[81] However, N348I and other connection mutations confer significant resistance to efavirenz and etravirine only when they co-evolve with NNRTI pocket mutations.[21,22]

The primary mechanism of resistance to AZT is by excision,[82,83] and RT develops EEMs that facilitate binding of ATP as excision agent.[17] The enhanced rigidity of the palm subdomain adjacent to the T215Y-mutated side chain may enhance ATP binding (Fig. 5); this putative effect may be analogous to the impact of L210W[84] accompanying the T215Y/F mutation to enhance ATP binding. This demonstrates that distal connection subdomain mutations contribute to drug resistance by altering dynamic properties of RT and the residue–residue contact network.

CONCLUSIONS

A molecular level understanding of the function of HIV-1 RT and the mechanism of NNRTI inhibition and resistance requires knowledge of both structure and dynamics. Hence, all-atom MD simulations of the RT-DNA, RT-DNA-NVP, and N348I/T369I mutant RT-DNA-NVP complexes were carried out. PCA analyses captured multiple key collective global motions of RT, and PSN analyses captured the subtle changes in the interaction network resulting from side-chain rearrangements of RT in the three different complexes. Global motions obtained from PCA-based essential dynamics analyses revealed heterogeneity in their motion that is different in different complexes of RT. NVP binding alters the global motions of RT and also restricts RT from sampling diverse conformations. Furthermore, PSN analyses carried on MD trajectory revealed a long-range

allosteric network in RT are also altered upon NVP binding or in the presence of connection subdomain mutations N348I/T369I. Insights gained from the present study provides a structural basis for understanding the resistance mechanism associated with the connection mutations, and the dynamic of coupling between the different subdomains of HIV-1 RT.

ACKNOWLEDGMENTS

MD simulations reported in this work have been partly performed at the BioMaPS High Performance Computing Center at Rutgers University. Authors thank Joseph D. Bauman, Disha Patel, and Steve Tuske for their insightful discussions and comments, Emilio Gallicchio and Nahren Manuel for their comments on the manuscript, and Matthew Miller and the BioMaPS technical support team for their assistance with computational hardware and software set up. This work was supported by US National Institutes of Health Awards AI 087201 to KD and AI 27690 (MERIT) to EA.

REFERENCES

1. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 1992;256:1783–1790.
2. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Jr., Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, Hizi A, Hughes SH, Arnold E. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. Proc Natl Acad Sci USA 1993;90: 6320–6324.
3. Steitz TA. DNA polymerases: structural diversity and common mechanisms. J Biol Chem 1999;274:17395–17398.
4. Ivetac A, McCammon JA. Elucidating the inhibition mechanism of HIV-1 non-nucleoside reverse transcriptase inhibitors through multicopy molecular dynamics simulations. J Mol Biol 2009;388:644–658.
5. Wright DW, Sadiq SK, De Fabritiis G, Coveney PV. Thumbs down for HIV: domain level rearrangements do occur in the NNRTI-bound HIV-1 reverse transcriptase. J Am Chem Soc 2012;134: 12885–12888.
6. Zhou Z, Madrid M, Evanseck JD, Madura JD. Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. J Am Chem Soc 2005;127: 17253–17260.
7. Monroe JI, El-Nahal WG, Shirts MR. Investigating the mutation resistance of non-nucleoside inhibitors of HIV-RT using multiple microsecond atomistic simulations. Proteins 2013; doi: 10.1002/prot.24346.
8. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. J Mol Biol 1999; 285:1023–1037.
9. Temiz NA, Bahar I. Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. Proteins 2002;49: 61–70.
10. Seckler JM, Leioatts N, Miao H, Grossfield A. The interplay of structure and dynamics: insights from a survey of HIV-1 reverse transcriptase crystal structures. Proteins 2013; doi: 10.1002/prot.24325.
11. Seckler JM, Howard KJ, Barkley MD, Wintrode PL. Solution structural dynamics of HIV-1 reverse transcriptase heterodimer. Biochemistry 2009;48:7646–7655.
12. Seckler JM, Barkley MD, Wintrode PL. Allosteric suppression of HIV-1 reverse transcriptase structural dynamics upon inhibitor binding. Biophys J 2011;100:144–153.
13. Abbondanzieri EA, Bokinsky G, Rausch JW, Zhang JX, Le Grice SF, Zhuang X. Dynamic binding orientations direct activity of HIV reverse transcriptase. Nature 2008;453:184–189.
14. Liu S, Abbondanzieri EA, Rausch JW, Le Grice SF, Zhuang X. Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. Science 2008;322:1092–1097.
15. Das, K, Arnold, E. HIV-1 reverse transcriptase and antiviral drug resistance. Part 1. Curr Opin Virol 2013;3:111–118.
16. Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, Arnold E. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. J Mol Biol 2009;385:693–713.
17. Tu X, Das K, Han Q, Bauman JD, Clark AD, Jr., Hou X, Frenkel YV, Gaffney BL, Jones RA, Boyer PL, Hughes SH, Sarafianos SG, Arnold E. Structural basis of HIV-1 resistance to AZT by excision. Nat Struct Mol Biol 2010;17:1202–1209.
18. Das K, Arnold E. HIV-1 reverse transcriptase and antiviral drug resistance. Part 2. Curr Opin Virol 2013;3:119–128.
19. Nikolenko GN, Palmer S, Maldarelli F, Mellors JW, Coffin JM, Pathak VK. Mechanism for nucleoside analog-mediated abrogation of HIV-1 replication: balance between RNase H activity and nucleotide excision. Proc Natl Acad Sci USA 2005;102:2093–2098.
20. Brehm JH, Koontz D, Meteer JD, Pathak V, Sluis-Cremer N, Mellors JW. Selection of mutations in the connection and RNase H domains of human immunodeficiency virus type 1 reverse transcriptase that increase resistance to 3′-azido-3′-dideoxythymidine. J Virol 2007;81: 7852–7859.
21. Yap SH, Sheen CW, Fahey J, Zanin M, Tyssen D, Lima VD, Wynhoven B, Kuiper M, Sluis-Cremer N, Harrigan PR, Tachedjian G. N348I in the connection domain of HIV-1 reverse transcriptase confers zidovudine and nevirapine resistance. PLoS Med 2007;4: e335.
22. Gupta S, Fransen S, Paxinos EE, Stawiski E, Huang W, Petropoulos CJ. Combinations of mutations in the connection domain of human immunodeficiency virus type 1 reverse transcriptase: assessing the impact on nucleoside and nonnucleoside reverse transcriptase inhibitor resistance. Antimicrob Agents Chemother 2010;54:1973–1980.
23. Nikolenko GN, Delviks-Frankenberry KA, Pathak VK. A novel molecular mechanism of dual resistance to nucleoside and nonnucleoside reverse transcriptase inhibitors. J Virol 2010;84:5238–5249.
24. Biondi MJ, Beilhartz GL, McCormick S, Gotte M. N348I in HIV-1 reverse transcriptase can counteract the nevirapine-mediated bias toward RNase H cleavage during plus-strand initiation. J Biol Chem 2010;285:26966–26975.
25. Schuckmann MM, Marchand B, Hachiya A, Kodama EN, Kirby KA, Singh K, Sarafianos SG. The N348I mutation at the connection subdomain of HIV-1 reverse transcriptase decreases binding to nevirapine. J Biol Chem 2010;285:38700–38709.
26. Das K, Martinez SE, Bauman JD, Arnold E. HIV-1 reverse transcriptase complex with DNA and nevirapine reveals non-nucleoside inhibition mechanism. Nat Struct Mol Biol 2012;19:253–259.
27. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 1998;282:1669–1675.
28. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997;18:2714–2723.
29. Sarafianos SG, Das K, Tantillo C, Clark AD, Jr., Ding J, Whitcomb JM, Boyer PL, Hughes SH, Arnold E. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. EMBO J 2001;20:1449–1461.

30. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. J Comput Chem 2005;26:1701–1718.
31. Das K, Sarafianos SG, Clark AD Jr, Boyer PL, Hughes SH, Arnold E. Crystal structures of clinically relevant Lys103Asn/Tyr181Cys double mutant HIV-1 reverse transcriptase in complexes with ATP and non-nucleoside inhibitor HBY 097. J Mol Biol 2007;5:77–89.
32. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–1174.
33. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 2006;25:247–260.
34. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. J Comput Chem 2002;23:1623–1641.
35. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera: a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–1612.
36. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 2003;24:1999–2012.
37. Jorgensen WL, Chandrasekhar J, Madura JD. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983; 79:926–935.
38. Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR. Molecular- dynamics with coupling to an external bath. J Chem Phys 1984;81:4368–4369.
39. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 1981;52:7182–7190.
40. Darden T, York D, Pedersen L. Particle mesh Ewald: an N log(N) method for Ewald Sums in large systems. J Chem Phys 1993;98: 10089–10092.
41. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997; 18:1463–1472.
42. Amadei A, Linssen AB, Berendsen HJ. Essential dynamics of proteins. Proteins 1993;17:412–425.
43. Poornam GP, Matsumoto A, Ishida H, Hayward S. A method for the analysis of domain movements in large biomolecular complexes. Proteins 2009;76:201–212.
44. Grant BJ, Rodrigues AP, ElSawy KM, McCammon JA, Caves LS. Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics 2006;22:2695–2696.
45. Kannan N, Vishveshwara S. Identification of side-chain clusters in protein structures by a graph spectral method. J Mol Biol 1999;292: 441–464.
46. Dijkstra EW. A note on two problems in connexion with graphs. Numerische Mathematik 1959;1:269–271.
47. Seeber M, Felline A, Raimondi F, Muff S, Friedman R, Rao F, Caflisch A, Fanelli F. Wordom: a user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. J Comput Chem. 2011;32:1183–1194.
48. Pasi M, Tiberti M, Arrigoni A, Papaleo E. xPyder: a PyMOL plugin to analyze coupled residues and their networks in protein structures. J Chem Inf Model 2012;52:1865–1874.
49. Palla G, Derenyi I, Farkas I, Vicsek T. Uncovering the overlapping community structure of complex networks in nature and society. Nature 2005;435:814–818.
50. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 2011;27:431–432.
51. Bec G, Meyer B, Gerard MA, Steger J, Fauster K, Wolff P, Burnouf DY, Micura R, Dumas P, Ennifar E. Thermodynamics of HIV-1 reverse transcriptase in action elucidates the mechanism of action of non-nucleoside inhibitors. J Am Chem Soc 2013; 135: 9743–9752.
52. Hess B. Convergence of sampling in protein simulations. Phys Rev E 2002;65:1–10.
53. Cramer CJ. Essentials of computational chemistry: theories and models. New York: John Wiley & Sons; 2004. p 93.
54. Luo G, Wang M, Konigsberg WH, Xie XS. Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. Proc Natl Acad Sci USA 2007;104:12610–12615.
55. Palaniappan C, Fay PJ, Bambara RA. Nevirapine alters the cleavage specificity of ribonuclease H of human immunodeficiency virus 1 reverse transcriptase. J Biol Chem 1995;270:4861–4869.
56. Chung S, Miller JT, Lapkouski M, Tian L, Yang W, Le Grice SF. Examining the role of the HIV-1 reverse transcriptase p51 subunit in positioning and hydrolysis of RNA/DNA hybrids. J Biol Chem 2013;288:16177–16184.
57. Batista PR, Costa MG, Pascutti PG, Bisch PM, de Souza W. High temperatures enhance cooperative motions between CBM and catalytic domains of a thermostable cellulase: mechanism insights from essential dynamics. Phys Chem Chem Phys 2011;13:13709–13720.
58. Altis A, Otten M, Nguyen PH, Hegger R, Stock G. Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis. J Chem Phys 2008;28:245102.
59. Wolynes PG, Onuchic JN, Thirumalai D. Navigating the folding routes. Science 1995;267:1619–1620.
60. Davis BC, Thorpe IF. Thumb inhibitor binding eliminates functionally important dynamics in the hepatitis C virus RNA polymerase. Proteins 2013;81:40–52.
61. Boehr DD, McElheny D, Dyson HJ, Wright PE. Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. Proc Natl Acad Sci USA 2010;107:1373–1378.
62. Fraser JS, Clarkson MW, Degnan SC, Erion R, Kern D, Alber T. Hidden alternative structures of proline isomerase essential for catalysis. Nature 2009;462:669–673.
63. Schmidtke P, Bidon-Chanal A, Luque FJ, Barril X. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. Bioinformatics 2011;27:3276–3285.
64. Popovych N, Sun S, Ebright RH, Kalodimos CG. Dynamically driven protein allostery. Nat Struct Mol Biol 2006;13:831–838.
65. Tsai CJ, del Sol A, Nussinov R. Allostery: absence of a change in shape does not imply that allostery is not at play. J Mol Biol 2008;378:1–11.
66. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: a plausible model. J Mol Biol 1965;12:88–118.
67. Koshland DE, Jr., Nemethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 1966;5:365–385.
68. Weber G. Ligand binding and internal equilibria in proteins. Biochemistry 1972;11:864–878.
69. Ma B, Kumar S, Tsai CJ, Nussinov R. Folding funnels and binding mechanisms. Protein Eng 1999;12:713–720.
70. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. Science 2008;320:1471–1475.
71. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Proteins 2004;57:433–443.
72. Daily MD, Gray JJ. Local motions in a benchmark of allosteric proteins. Proteins 2007;67:385–399.
73. Brinda KV, Vishveshwara S. A network representation of protein structures: implications for protein stability. Biophys J 2005;89: 4159–4170.
74. Brinda KV, Vishveshwara S. Oligomeric protein structure networks: insights into protein–protein interactions. BMC Bioinformatics 2005;6:296.
75. Tachedjian G, Orlova M, Sarafianos SG, Arnold E, Goff SP. Nonnucleoside reverse transcriptase inhibitors are chemical enhancers of

dimerization of the HIV type 1 reverse transcriptase. Proc Natl Acad Sci USA 2001;98:7188–7193.

76. del Sol A, O'Meara P. Small-world network approach to identify key residues in protein–protein interaction. Proteins 2005;58:672–682.
77. Tachedjian G, Aronson HE, de los Santos M, Seehra J, McCoy JM, Goff SP. Role of residues in the tryptophan repeat motif for HIV-1 reverse transcriptase dimerization. J Mol Biol 2003;326:381–396.
78. Gopalakrishnan V, Peliska JA, Benkovic SJ. Human immunodeficiency virus type 1 reverse transcriptase: spatial and temporal relationship between the polymerase and RNase H activities. Proc Natl Acad Sci USA 1992;15:10763–10767.
79. Balzarini J, Auwerx J, Rodriguez-Barrios F, Chedad A, Farkas V, Ceccherini-Silberstein F, Garcia-Aparicio C, Velazquez S, De Clercq E, Perno CF, Camarasa MJ, Gago F. The amino acid Asn136 in HIV-1 reverse transcriptase (RT) maintains efficient association of both RT subunits and enables the rational design of novel RT inhibitors. Mol Pharmacol 2005;68:49–60.
80. Ehteshami M, Beilhartz GL, Scarth BJ, Tchesnokov EP, McCormick S, Wynhoven B, Harrigan PR, Gotte M. Connection domain mutations N348I and A360V in HIV-1 reverse transcriptase enhance resistance to 3′-azido-3′-deoxythymidine through both RNase H-dependent and -independent mechanisms. J Biol Chem 2008;283: 22222–22232.
81. Das K, Clark AD, Jr., Lewi PJ, Heeres J, De Jonge MR, Koymans LM, Vinkers HM, Daeyaert F, Ludovici DW, Kukla MJ, De Corte B, Kavash RW, Ho CY, Ye H, Lichtenstein MA, Andries K, Pauwels R, De Bethune MP, Boyer PL, Clark P, Hughes SH, Janssen PA, Arnold E. Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. J Med Chem 2004;47:2550–2560.
82. Arion D, Kaushik N, McCormick S, Borkow G, Parniak MA. Phenotypic mechanism of HIV-1 resistance to 3′-azido-3′-deoxythymidine (AZT): increased polymerization processivity and enhanced sensitivity to pyrophosphate of the mutant viral reverse transcriptase. Biochemistry 1998;37:15908–15917.
83. Meyer PR, Matsuura SE, So AG, Scott WA. Unblocking of chain-terminated primer by HIV-1 reverse transcriptase through a nucleotide-dependent mechanism. Proc Natl Acad Sci USA 1998;95: 13471–13476.
84. Boyer PL, Sarafianos SG, Arnold E, Hughes SH. Selective excision of AZTMP by drug-resistant human immunodeficiency virus reverse transcriptase. J Virol 2001;75:4832–4842.